DOCUMENT RESUME

ED 268 145 TM 860 070

AUTHOR Marso, Ronald N.
TITLE Testing Practices and Test Item Preferences of Classroom Teachers.
PUB DATE Oct 85
NOTE 21p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (Chicago, IL, October 17-19, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Constructed Response; *Educational Testing; Elementary Secondary Education; Higher Education; Instructional Program Divisions; Multiple Choice Tests; Questionnaires; Surveys; *Teacher Attitudes; *Teacher Made Tests; Teaching Experience; *Test Construction; *Test Format; Test Items; *Test Use

ABSTRACT

A questionnaire concerning the use of teacher made tests was completed by 123 public school teachers in Ohio. Five testing practices were examined: (1) number of teacher made tests given during a course or school year; (2) types of test items most commonly used; (3) sources used in obtaining test items; (4) information used in assigning grades; and (5) relative usefulness of the test item types. The responses were analyzed by grade level (K-4, 5-8, or 9-12) and by years of teaching experience (1-8, 9-13, or 14 or more). Results indicated that a large number of teacher made tests were administered; the mean frequency was 99.5 tests given during a typical school year. The mean frequency for a single course was 23.2, and teachers with 9 to 13 years of experience gave the greatest number of tests. The most common test item type was completion, followed by problem exercises, matching, multiple choice, essay, and true-false. Teachers' ratings of item usefulness favored matching items, followed by multiple choice, completion, true-false, problem exercises, and essays. Differences were noted among teachers' grade levels for perceived usefulness of test item types. Self-constructed items were the most important source of test items, used by 96 percent of the respondents. Other sources included textbook manuals (77 percent), workbooks (58 percent), and old tests (49 percent). Test scores were used by 97 percent of the teachers in assigning grades. Other criteria included daily work, class participation, and student projects. (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED268145

Testing Practices and Test Item Preferences of Classroom Teachers

Midwestern Educational Research Association
1985 Annual Meeting

Chicago, October 17-19

Ronald N. Marso
Bowling Green State University
Bowling Green, Ohio 43403

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. N. Marso

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 860 070

AUTHOR

MARSO, RONALD N. Address: College of Education, Bowling Green University, Bowling Green, Ohio 43403. Title: Professor, Educational Foundations and Inquiry. Degrees: B.A. Dakota State, M.A. Adams State, Ed.D. University of Nebraska. Specialization: Educational Psychology and Measurement. Telephone: 1-419-372-0151.

Abstract

A total of 123 teachers responded to a survey of classroom testing practices encompassing questions within five categories: number of teacher-made tests given during a typical course and school year; types of test items most commonly used; sources used in obtaining test items; information used in assigning grades; and relative usefulness of the common test item types. Although significant differences were identified within each of these five categories when responses were analyzed by grade and teaching experience levels, these differences tended to be relatively minor variations within rather common testing practices. These teachers reported giving large numbers of teacher-made tests; having a preference for self-constructed test questions; viewing matching, multiple-choice, and completion type test items as being most useful; using test scores as a major factor in assigning grades; and most frequently using completion, problem exercises, and multiple-choice items, respectively.

Testing Practices and Test Item Preferences of Classroom Teachers

Although extensive literature exists advising classroom teachers about testing and evaluation practices, it appears that relatively little research has been conducted on teacher-made tests and related testing practices in the public schools (Coffman, 1971). Fleming and Chambers (1983) state that these tests occupy considerable amounts of classroom time and effort but are a little researched part of the drama of testing in classrooms. Further, Dwyer (1982) maintains that the advice to teachers regarding teacher constructed tests tends to reflect a consensus of professional judgements rather than data obtained from research. Little is known about what test item types are preferred by classroom teachers; where teachers obtain their items; how frequently they give tests; their perceptions of the usefulness of various item types; or the relative contribution of test scores to course grades.

Measurement specialists and those preparing classroom teachers continue to emphasize the importance of test item selection, of the relative frequency of tests, of test development practices leading to high levels of test reliability and validity, and how these factors impact upon learning outcomes. This, however, appears to occur in the absence of knowledge of typical teacher-made test development and testing practices.

Problem

The general purpose of this study was to investigate the classroom testing practices of teachers in the public schools. The following five hypotheses were stated and tested to guide this investigation:

1. Teachers will report frequent use of teacher-made tests during an academic school year.
   1.a. Reported frequencies will vary among grade levels.
   1.b. Reported frequencies will vary among years of teaching experience levels.
2. Teachers will report preferences in the use of the common test item types.
   2.a. Reported preferences will vary among grade levels.
   2.b. Reported preferences will vary among years of teaching experience levels.
3. Teachers will report differing perceptions of the usefulness of common test item types.
   3.a. Reported perceptions will vary among grade levels.
   3.b. Reported perceptions will vary among years of teaching experience levels.
4. Teachers will report that self-constructed test items are their most important source of items in preparing tests.
5. Teachers will report that test scores are the most important factor in assigning course grades.

Method

A questionnaire was constructed and sent to area classroom teachers registered as part-time graduate students during the fall semester of 1983 at Bowling Green State University. These teachers resided within a fifteen county area in Northwest Ohio surrounding the University. The questionnaires were mailed with a cover letter and a postage paid response envelope. The cover letter solicited teacher

cooperation in reporting daily testing practices to assist the University in more realistically preparing new teachers. A total of 123 useable questionnaires was obtained from currently active classroom teachers from a potential computer generated pool of approximately 298 names and addresses.

Information was requested about five categories of testing practices closely related to the five stated hypotheses. These categories were: the number of tests given in a typical school year by subject area; the frequency rating of types of test questions used; a checklist of the usefulness of each test question type; a rating of the relative importance of common sources of test questions; and a rating of the relative importance of types of student information used in assigning course grades. Each information category included an "other" option in the event that items within a category might not seem inclusive to a respondent. Each respondent also reported his or her years of teaching experience, and grade or subject assignment.

Results

The 123 teacher responses to the survey were analyzed by grade level, K-4, 5-8, and 9-12 grades, and by years of teaching experience, 1-8, 9-13, and 14 or more years. The numbers of teacher respondents within these categories were 30, 46, 47, and 46, 45, and 32, respectively. Tallies within each information category were counted, percentages of respondents computed, and mean ratings computed and ranked for those categories soliciting a rating response. The results of this process are reported by hypothesis tested and questionnaire category in the following paragraphs and in tabular form on tables one through three.

1) Hypothesis one, frequency of tests: The respondents reported a mean frequency of 99.5 tests being given during a typical school year. These responses tallied by grade level resulted in the following mean frequencies: K-4 grades = 97.9, 5-8 grades = 139.0, and 9-12 grades = 70.1. These academic year testing frequencies by grade level groups vary sufficiently to reject a goodness of fit null hypothesis at the $p < .001$ level (X^2 of 23.50, df = 2) thus supporting hypothesis 1.a. The mean frequency of tests given during a single course was 23.2 for the total group, and means of 26.1, 20.4, and 19.5 were reported respectively for the 5-8, K-4 and 9-12 grade levels. Most teachers reported a per course test frequency within the 15 to 20 range with the exception of spelling classes most commonly following a weekly testing pattern. Teachers in the intermediate years of experience category reported giving the most tests per academic year with means of 89.8, 127.2, and 105.7 for 1-8, 9-13, and 14 or more years of teaching experience, respectively. These reported frequencies vary sufficiently to reject a goodness of fit null hypothesis at the $p < .05$ level (X^2 of 6.85, df = 2), thus supporting hypothesis 1.b.

2) Hypothesis two, test question types used: The following test item type mean frequencies of use on a five point scale were computed from the questionnaire responses from highest to lowest rank order: completion type 3.27, problem exercises 3.13, matching 3.06, multiple-choice 3.04, essay type 2.78, and true-false 2.49. These rankings were relatively stable when the data was analyzed by grade level, Kendall's coefficient of concordance of .90 ($p < .05$), and within years of teaching experience, Kendall's coefficient of concordance of .82 ($p < .05$). The completion item type was consistently ranked highest,

the true-false item type lowest, and essay item type second from the lowest.

It is interesting to note, however, that the ranking of item types on the basis of frequency of use had little agreement with the ranking of item types by perceived usefulness for this sample of teachers. The Spearman rank correlation estimate between these two sets of ranks yielded a coefficient of only .26. This would suggest that availability of test item types, construction ability, or other unknown factors influenced actual test item use as well as teacher perception of item usefulness.

The analysis of teacher responses to the set of item types did support hypothesis two which predicted a difference in reported frequency in the use of common test item types. An independence of response pattern test resulted in a X^2 value of 80.64, df = 20, significant at the $p < .001$ level of confidence.

When this data was analyzed by teaching experience and grade level of the teachers, no significant differences were found for years of teaching experience while several differences were identified among the comparisons by grade levels of the respondents. The differences in response patterns to the total set of test item types by grade level were significant at the $p < .05$ level (X^2 value of 9.32, df = 2). Analysis of teacher responses to specific test item types by grade level revealed differences in reported use for the completion item type, X^2 of 25.53, df = 8, $p < .05$; for matching type, X^2 of 15.77, df = 8, $p < .05$, and a borderline difference for multiple-choice item, $X^2 = 15.23$, df = 8, $p < .06$. Each of these item types was reported to be used less frequently in the lower grade levels compared to the

upper grades. The mean frequency rating and ranks for the items by the total sample and by age and grade levels are reported on Tables 1 and 2.

3) Hypothesis three, test item usefulness: The respondents reported item type usefulness by checking one or more of the following six positive statements for each item type: lends itself well to subject; easy to prepare and score; accepted and familiar to me; helps reach learner goals; accurate and fair; and available in text manual. An analysis of these checks resulted in the following rankings for the total group of respondents from highest to lowest number of statements checked: matching, multiple-choice, completion, true-false, problem exercises, and essay item types. These rankings remained relatively stable within both the teacher experience and grade level classifications with matching and multiple-choice items given top rankings and essay items the bottom ranking.

The tallies to the individual usefulness statements indicate that: matching, multiple-choice, and completion respectively were perceived as being most fair; essay and problem exercises were reported as difficult to use; matching, multiple-choice, and true-false items respectively were viewed as easiest to prepare and score; matching, completion, and multiple-choice respectively were considered most acceptable to the teachers; and matching, multiple-choice, and completion item types respectively were reported as most adaptable to their subject matter.

Responses to the total set of item types suggested differing perceptions of the usefulness of the various types of test items, thus

supporting hypothesis three. The independence of response pattern test for the total sample yielded a X^2 of 78.37, df = 5, $p < .001$.

Analysis of the response patterns within teacher grade levels produced significant differences in the reported usefulness for each item type with the exception of the problem exercises, thus lending support to hypothesis 3.a. The 9-12 grade level teachers checked relatively more positive statements for the essay (X^2 of 10.71, df = 2, $p < .01$) and for the true-false item types (X^2 of 13.79, df = 2, $p < .01$). The 5-8 grade level teachers checked relatively more positive statements for the completion (X^2 of 10.18, df = 2, $p < .01$), for the matching (X^2 of 6.66, df = 2, $p < .05$), and for the multiple-choice type tests items (X^2 of 7.75, df = 2, $p < .05$).

Comparisons among the teaching experience levels did not result in significant differences with the exception of responses to the multiple-choice item type. On this item type the teachers with 1-8 years of experience checked relatively more positive statements than did the other teachers (X^2 of 6.40, df = 2, $p < .05$). Thus little support was found for hypothesis 3.b.

4) Hypothesis four, sources of test questions: Responses to the importance of sources for obtaining test questions on a four point scale resulted in mean ratings as follows: self-constructed 3.60, workbooks 2.46, textbook manuals 2.39, and old tests 2.11. The following percentages of teachers reported that they obtained test items from these sources: 96% self-constructed, 77% textbook manuals, 58% workbooks, and 49% old tests.

An analysis of the responses by the total sample of teachers on this set of sources for obtaining test items resulted in a X^2 value of

154.48, df = 9, $p < .001$ indicating a difference in response patterns among the sources of test items. Self construction was reported as the most important source and was rated significantly more important by the teachers than the second highest source of workbooks at the $p < .001$ level of confidence (X^2 of 79.66, df = 3) thus providing strong support for hypothesis four. The self-constructed items were also reported as most often used by the teachers with a goodness of fit null hypothesis rejected at the $p < .001$ level (X^2 of 23.33, df = 3). The percentage of teachers reporting use, mean rating, and relative ranking of the identified sources are reported on Tables 1 and 2.

When making comparisons of responses to the test item sources by grade level and years of teaching experiences, significant differences were identified only with workbook and self-constructed sources. Workbooks were reported as a more frequent source of test items for K-4 grade teachers than other teachers (X^2 of 6.82, df = 2, $p < .05$) and were reported as a more important source by 5-8 grade level teachers compared to the other teachers (X^2 of 16.12, df = 6, $p < .05$). The self-constructed source of test items was reported as a more important source by the secondary teachers as compared to the other teachers (X^2 of 19.21, df = 4, $p < .001$). Similarly, the rank orderings of the sources of test items were relatively consistent within these two groups. Kendall's coefficient of concordance was .64 among the grade level rankings and was .73 among the teaching experience level rankings.

5) Hypothesis five, importance of tests for grades: The responses within this information category indicated that 97% of all the respondents

used test scores in assigning grades. These responses also suggested that daily work (97%), participation in class (87%), and student projects (76%) also commonly contributed to the final course grade. These relative percentages were quite consistent across groups with no significant differences among the reported frequencies.

In terms of reported relative importance of these sources of information in assigning course grades, the mean rating of importance on a four point scale from highest to lowest follows: quizzes and tests 3.19, daily work exercises and homework 3.04, special projects and term papers 2.61, and class participation 2.40. This rank order was identical for the 5-8 and 9-12 grade levels among the years of teaching experience. The K-4 group of teachers ranked daily work, tests, class participation, and student projects from most to least important in grading. The Kendall coefficient of concordanc was .88 among the grade level rankings and of course 1.00 among the identical rankings by years of teaching experience.

Teacher response patterns to this total category resulted in a significant independence test (X^2 of 91.37, df = 9, $p < .001$) indicating a differing response to items contributing to the assignment of course grades. Tests and quizzes were reported to be the most important factor within all comparison groups with the exception of the single K-4 teacher grouping which rated daily work as more important. This group rated daily work significantly higher in importance than tests in assigning course grades (X^2 of 8.35, df = 2, $p < .02$). The 5-8 grade level teachers rated tests highest and significantly more important than the second ranked daily work source of information (X^2

of 6.30, df = 2, $p < .05$). The 9-12 grade teachers also rated tests as most important but not significantly more so than the second ranked category of daily work information (X^2 of 2.91, df = 2, NS), although tests were rated significantly greater than the third ranked category of term projects (X^2 of 6.01, df = 2, $p < .05$).

The total teacher response pattern provided support for hypothesis five stating that tests and quizzes are the most important source of information in assigning course grades. Comparisons by grade level of instruction and years of teaching experience, however, indicated this was true in only six of the seven groupings. It must be noted that K-4 level teachers in this sample rated daily work significantly more important than tests scores in assigning course grades while 9-12 grade level teachers reported tests as most important but not significantly more important than daily work performance. Another difference approaching significance was noted within this response category. The secondary teachers rated term projects relatively higher than the teachers at other grade levels (X^2 of 11.95, df = 6, $p < .07$).

Conclusions and Discussion

Hypothesis one: The teachers reported a high mean frequency of 99.5 classroom tests being given during a typical school year. Additionally, it was found that test frequencies varied significantly among years of teaching experience and grade levels of instruction. Highest frequencies of tests were reported by teachers with 9-13 years of experience and those teaching in the 5-8 grades. This hypothesis was accepted.

Hypothesis two: The respondents reported a varied preference for using the common test item types which supported this general hypothesis. Reported item use frequencies also varied by grade level but not significantly by years of teaching experience. Thus hypothesis 2.a. but not 2.b. was supported by the obtained data. In general the teachers reported use of a variety of test items with item types ranked from highest to lowest frequency: completion, problem exercises, matching, multiple-choice, essay, and true-false. Variance found by grade level appeared to be due primarily to less frequent use of matching, multiple-choice, and completion items in the lower grades.

Hypothesis three: The teachers in this sample reported differing levels of usefulness for the common test item types. When these responses were analyzed by grade level and years of teaching experience, comparisons by grade level resulted in a number of significant differenc s while comparisons by experiences resulted in only one significant difference. Thus the general hypothesis and 3.a. but not 3.b. were supported by the data obtained.

Generally matching, multiple-choice, completion, true-false, problem exercises, and essay type items were ranked from most to least useful. The secondary teachers reported essay and true-false items as being more useful than the other teachers. The middle grade teachers checked more positive statements than the other teachers for completion, matching, and multiple-choice items. The single significant difference noted among the teaching experience level comparisons indicated that the teachers with fewer years of experience viewed the multiple-choice type item more positively than did the more experienced teachers.

Hypothesis four: The teachers rated the self-constructed source as significantly more important than any other source of test items in preparing classroom tests. Additionally, this source of test items was ranked highest in importance in all comparisons by grade or teaching experience levels and was reported as the most frequently used source, thus providing clear support for this hypothesis.

The general rank ordering for the source of test items from high to low was self-constructed, workbooks, textbook manuals, and old tests. The only differences noted within this category were a relatively higher rating of the workbook source by K-4 teachers and a relatively higher rating of self-constructed items by the secondary teachers.

Hypothesis five: The teachers rated test scores as the most important factor in assigning student grades which supports this hypothesis. The relative rankings within this category were very consistent among the comparison groups with the single exception of the K-4 grade teachers, who ranked daily work above test scores in assigning grades, although this difference was not significant statistically. Similarly, although the secondary teachers ranked test scores as highest, test scores were not rated significantly higher than daily work. Thus the data supports this hypothesis with the caution that daily work was rated as highly as test scores in contributing to course grades by the K-4 and 9-12 grade level teachers in this sample.

When asked to check the usefulness of various item types such as ease of use, fairness, and helping learners reach goals, teacher responses resulted in the following high to low rank order for the

items: matching, multiple-choice, completion, true-false, problem exercises, and essay. The matching and multiple-choice item types were reported as most acceptable to the teachers; these two item types along with completion items were viewed as readily adaptable to subject content; and the matching, multiple-choice, and completion items were perceived as most accurate and fair to students.

Approximately one-half of the teachers reported changes of emphasis in the identified testing patterns in recent years as reported in Table 3. This pattern of change appeared to be stable over both grade level and years of teaching experience. Although the change in emphasis reported was not always in the same direction, increased emphasis was apparent for the following testing practices listed in order of mean rank: emphasis on student performance, basic skill measurement, frequency of tests, and use of essay questions.

In summation, this teacher sample reported giving large numbers of tests during an academic school year; having a preference for self-constructed test items; viewing test scores as a major factor in assigning course grades; using completion, problem exercises, and matching items most frequently; and perceiving matching, multiple-choice and completion type test items as being most useful.

9/3

References

Coffman, W. E. (1971). Essay examinations. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 271-302). Washington, D.C.: American Council on Education.

Dwyer, C. A. (1982). Achievement testing. In H.E. Mitzel (Ed.), Encyclopedia of educational research (4th ed., Vol. 1, pp. 13-22). New York: The Free Press.

Fleming, M. & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. New directions for testing and measurement. 19, 29-38.

TABLE 1
Responses by Grade Levels
Information Categories II, III, IV, and V

II. Test Question Types Used: Responses on five point scale from always to never used.

Grade Levels:	K-4		5-8		9-12		Total	
Item Types	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
True-false	2.07	6	2.51	6	2.73	6	2.49	6
Matching	2.71	4	3.13	2	3.20	2	3.06	3
Completion	3.00	1	3.26	1	3.47	1	3.27	1
Multiple Choice	2.85	3	3.02	4	3.19	3	3.04	4
Essay	2.60	5	2.69	5	3.13	5	2.78	5
Problem exercises	2.90	2	3.12	3	3.18	4	3.13	2

III. Test Question Type Usefulness: Percentage of checks to total possible and relative ranking of numbers of usefulness statements checked for item types.

Grade Levels:	K-4		5-8		9-12		Total	
Item Types	%	Rank	%	Rank	%	Rank	%	Rank
True-false	.23	5	.35	5	.38	3	.33	4
Matching	.44	2	.59	1	.44	1	.50	1
Completion	.42	3	.47	3	.37	4	.42	3
Multiple Choice	.45	1	.53	2	.41	2	.46	2
Essay	.16	6	.23	6	.29	5.5	.23	6
Problem exercises	.32	4	.35	4	.29	5.5	.32	5

IV. Sources of Test Items: Importance of source of items on a four point scale from most to minor importance and % reporting use.

Grade Levels:	K-4		5-8		9-12		Total		
Sources	% yes	Mean	% yes	Mean	% yes	Mean	% yes	Mean	Rank
Text manuals	.97	2.76	.77	2.29	.64	2.17	.77	2.39	3
Workbooks	.86	2.24	.60	2.52	.38	2.21	.58	2.46	2
Old tests	.48	2.14	.43	1.95	.55	2.22	.49	2.11	4
Self constructed	.97	3.30	.94	3.53	.98	3.85	.96	3.60	1

V. Contribution to Grades: Contributions to assignment of final grade on a four point scale from more to less important and % reporting use.

Grade Levels:	K-4		5-8		9-12		Total		
Sources	% yes	Mean	% yes	Mean	% yes	Mean	% yes	Mean	Rank
Class part.	.93	2.43	.87	2.31	.84	2.47	.87	2.40	4
Daily work	1.00	3.20	.96	2.93	.96	3.06	.97	3.04	2
Tests	.93	2.88	1.00	3.28	.96	3.26	.97	3.19	1
Term projects	.73	2.25	.90	2.54	.87	2.87	.76	2.61	3

TABLE 2
Responses by Years of Experience
Information Categories II, III, IV, and V

II. Test Question Types Used: Responses on five point scale from always to never used.

Years of Experience:	1-8		9-13		14+		Total	
Item Types	Mean	Rank	Mean	Rank	Mean	Rank	Mean	Rank
True-false	2.70	6	2.34	6	2.42	6	2.49	6
Matching	3.21	2	3.00	4	2.93	3	3.06	3
Completion	3.39	1	3.21	1	3.20	2	3.27	1
Multiple Choice	3.17	3	3.05	2	2.83	4	3.04	4
Essay	2.88	5	2.73	5	2.63	5	2.78	5
Problem exercises	3.05	4	3.03	3	3.38	1	3.13	2

III. Test Question Type Usefulness: Percentage of checks to total possible and relative ranking of numbers of usefulness statements checked for item types.

Grade Levels:	1-8		9-13		14		Total	
Item Types	%	Rank	%	Rank	%	Rank	%	Rank
True-false	.36	4	.29	5	.33	5	.33	4
Matching	.51	2	.48	1	.49	1	.50	1
Completion	.42	3	.42	3	.40	2	.42	3
Multiple Choice	.54	1	.43	2	.39	3.5	.46	2
Essay	.23	6	.22	6	.28	6	.23	6
Problem exercises	.27	5	.30	4	.39	3.5	.32	5

IV. Sources of Test Items: Importance of source of items on a four point scale from most to minor importance and % reporting use.

Grade Levels:	1-8		9-13		14+		Total		
Sources	% yes	Mean	% yes	Mean	% yes	Mean	% yes	Mean	Rank
Text manuals	.79	2.37	.82	2.33	.67	2.55	.77	2.39	3
Workbooks	.64	2.63	.53	2.24	.55	2.53	.58	2.46	2
Old tests	.55	2.48	.51	1.87	.36	1.86	.49	2.11	4
Self constructed	.98	3.55	.91	3.64	1.00	3.68	.96	3.60	1

V. Contribution to Grades: Contributions to assignment of final grade on a four point scale from more to less important and % reporting use.

Grade Levels:	1-8		9-13		14+		Total		
Sources	% yes	Mean	% yes	Mean	% yes	Mean	% yes	Mean	Rank
Class part.	.89	2.51	.86	2.38	.87	2.29	.87	2.40	4
Daily work	.96	3.09	1.00	2.98	.94	2.93	.97	3.04	2
Tests	.96	3.12	.98	3.37	.97	3.13	.97	3.19	1
Term projects	.77	2.83	.93	2.39	.73	2.68	.76	2.61	3

TABLE 3
Total Group Response Frequencies
Information Categories II, IV, V, and VI

II. Test Question Types Used:	Never Use N	Never Use %	Seldom Use N	Seldom Use %	Commonly Use N	Commonly Use %	Nearly Always N	Nearly Always %	Always Use N	Always Use %
1. True-false type questions	18	16	45	39	35	30	14	12	4	3
2. Matching type exercises	6	5	23	19	53	45	30	25	6	5
3. Completion type questions	4	3	14	12	57	47	37	31	9	7
4. Multiple-choice questions	8	7	27	22	44	38	26	22	11	9
5. Essay type questions	16	13	40	34	30	25	20	17	13	11
6. Problem solution exercises	9	9	23	22	36	35	16	16	19	18

IV. Sources of Test Items	Use this Source N	Use this Source %	Minor Importance N	Minor Importance %	Somewhat Important N	Somewhat Important %	Very Important N	Very Important %	Most Important N	Most Important %
1. Teacher's manuals for textbooks	95	77	16	17	35	37	36	38	8	8
2. From workbook exercises	71	58	8	11	26	37	33	46	4	6
3. Old tests and other teachers	60	49	16	27	24	40	18	30	2	3
4. Questions written by yourself	118	96	0	0	6	5	35	30	77	65

V. Contribution to Grades	Use in Grading N	Use in Grading %	Minor Importance N	Minor Importance %	Somewhat Important N	Somewhat Important %	Very Important N	Very Important %	Most Important N	Most Important %
1. Participation in class activities	107	87	16	15	40	37	43	40	8	7
2. Daily work exercises and homework	119	97	1	1	17	14	77	65	24	20
3. Quizzes and tests	119	97	0	0	14	12	70	59	35	29
4. Special projects/ term papers	94	76	8	9	31	33	44	47	11	12

VI. Changing Test Practices	Reporting Change N	Reporting Change %	Emphasis More N	Emphasis More %	Somewhat More N	Somewhat More %	Somewhat Less N	Somewhat Less %	Less N	Less %
1. Change in the numbers of tests given?	79	65	19	24	33	42	22	28	5	6
2. Change in use of essay questions?	57	47	14	25	19	33	16	28	8	14
3. Change in emphasis on grades?	63	51	12	19	18	29	30	48	3	5
4. Change in measurement of basic skills?	67	56	25	37	29	43	10	15	3	4
5. Availability of test questions in manuals, etc.	48	39	9	19	22	46	11	23	6	13
6. Change in emphasis on student performance?	76	62	34	45	30	40	9	12	3	4